A Location Privacy Protection Framework with Mobility Using Host Identity Protocol

AUTHOR(S):

MAEKAWA, Keiji

Master Thesis

A Location Privacy Protection Framework with Mobility Using Host Identity Protocol

Supervisor Professor Yasuo Okabe

Department of Intelligence Science and Technology
Graduate School of Informatics
Kyoto University

Keiji MAEKAWA

February 6, 2009

A Location Privacy Protection Framework with Mobility Using Host Identity Protocol

Keiji MAEKAWA

Abstract

Mobility is a key element of the future Internet. The location privacy problem is one of the problems involved in mobility. A great benefit of universally available access to the Internet might bring a risk such that a user's location is traceable by others.

Most of the mobility protocols define a mechanism of informing the correspondents of location change, in order to realize mobility. Therefore, the correspondents and eavesdroppers on the path will notice a movement and its destination.

The problem changes its situation, according to the node from which the location of a mobile node should be concealed. We classify the nodes into two types: correspondents or onlookers on the path. In addition, we often assume there are some trustworthy nodes on the path.

There are some existing researches on this problem, e.g. HIP Location Privacy Framework by Matos et al. and BLIND by Ylitalo and Nikander. They showed that it is possible to conceal location from a correspondents and a part of onlookers by introducing a trustworthy helper node, and especially when mobility is not needed at all, from all the onlookers, too.

In this research, we have proposed a new framework using Host Identity Protocol (HIP), and with it we showed that it is also possible to protect location privacy from all other nodes in IP communication with mobility. We take advantage of the notable feature of HIP that public keys are used as host identifiers, so that our framework gives a way to separate IDs for mobility from those for end-to-end communication. We constructed an extensional mobility management protocol of BLIND, and discussed a trade-off in terms of efficiency and operational cost.

Host Identity Protocolを用いた 移動通信のためのロケーションプライバシ保護フレームワーク

前川 慶司

内容梗概

モビリティ技術は今後のインターネットの発展を考える上で重要な位置を占める要素である．モビリティと深く関連する問題のひとつとして，ロケーションプライバシ問題がある．あらゆる場所からインターネットへアクセスが可能となる利便性の裏には，ユーザの位置を第三者に追跡されるリスクがつきまとう．多くの場合，モビリティプロトコルでは通信相手に自身の移動を知らせることによってモビリティを実現する．そのため通信相手や盗聴者が移動ノードの位置の変化およびその移動先を知ることになる．

この問題はユーザの位置情報を誰に対して秘匿するかによって状況が異なる．秘匿対象として通信相手と通信経路上の第三者の二種類を考え，さらに後者の一部に信頼できるノードを仮定する場合がある．

この問題に対する従来研究として Matos らによる HIP Location Privacy Framework や Ylitalo らによる BLIND などがあり，信頼できる補助ノードを導入することで通信相手や一部の盗聴者に対する位置の秘匿が可能であることや，モビリティを考慮しない状況に限れば，通信相手および通信経路上の全ノードに対する位置の秘匿が可能であることが知られている．

本研究において我々は Host Identity Protocol (HIP) を使った新たな手法を提案し，ネットワーク間の移動を伴う IP 通信においても，すべての対象に対するロケーションプライバシの保護が可能であることを示した．

我々の手法では，公開鍵がホストの識別子として使われるという HIP の特徴を活用し，移動用の ID と通信用の ID とを分離する．これに基づいて BLIND に対してモビリティ管理を行うための拡張プロトコルを構成し，モビリティとロケーションプライバシの両立に伴う通信効率や運用コストとのトレードオフについて考察した．

A Location Privacy Protection Framework with Mobility Using Host Identity Protocol

Contents

Chapter 1 Introduction

The Internet is now moving on to the next stage. Wireless infrastructures are getting more commonly available, and IPv6 network is actively promoted to replace the IPv4 network. Compared to the days when the current Internet was designed, circumstances are totally different.

Mobility is a key concept here. A flexible architecture is needed in a place where mobile nodes freely roam around networks and communicate with each other, dynamically changing its address. As well as mobility, security and privacy have become very important. It is desired that people are protected from gathering and analyzing sensitive user data.

The location privacy problem is a common problem among mobility protocols. In order to keep a session, a mobile node has to inform another node of the location change. This location updating message easily reveals the location of the mobile node to others. It is difficult to give a complete solution to this. So, there are several proposals to solve this problem, which offer partial solutions with a trustworthy helper node.

Our purpose is to integrate mobility and location privacy. We use Host Identity Protocol (HIP) as the core mobility protocol. HIP is designed for realizing mobility and multihoming IP environment in a secure way. The "identifier is public key" principle is natural and powerful in dealing with security and privacy. We take advantage of this property to construct a location privacy protection framework with mobility support in this paper.

The rest of this paper is organized as follows. First, some basic knowledge in IP mobility and location privacy is introduced in Chapter 2. In Chapter 3, two location privacy frameworks are explained, and compared in achievement. Then, we propose the design of a new location privacy framework in Chapter 4. Finally, we evaluate our framework in Chapter 5, and concludes in Chapter 6.

Chapter 2 Preliminaries

In this chapter, we introduce some basic concepts and terminology in IP mobility and related privacy issues, which take an important part in the later discussion.

We explain basics of IP mobility in 2.1, with illustrations of Mobile IP and LIN6. Location privacy is defined in 2.2. Host Identity Protocol is introduced in 2.3, with the properties on its mobility and location privacy.

2.1 Transparent IP Mobility

Mobile Node (MN) is a node which can roam over networks, and *Correspondent Node* (CN) is a node which corresponds with an MN. CN is possibly mobile node, too.

Some mobility protocols, such as Mobile IP[1][2], LIN6[3] and HIP[4], provide mobility support. With these protocols, an MN can change its location and the change is transparent to upper layers. End-point applications are totally independent of the roaming and do not have to manage location information. Therefore legacy applications also work without upgrade.

2.1.1 Identifier/Locator Split

In the current TCP/IP network, IP address has two roles: end-point identifier and locator of a host.

Applications use end-point identifiers so as to distinguish hosts and designate particular correspondents. On the other hand, locators represent a topological location in the network and are used as routing information in packet delivery.

One of shortcomings in the current TCP/IP architecture is the fact that it does not support host mobility. IP address is locator defined in the network layer, however, applications regard it as an end-point identifier, and assume that IP address of a correspondent remains same during a whole session. Therefore, a session will be broken when the host changes its IP address, unless applications take a special care for that.

If locators and identifiers are separated clearly, it becomes much easier to realize transparent mobility. For example, Mobile IP defines *Home Address* (HoA) as an identifier and *Care-of Address* (CoA) as a locator. Home Address

is the IP address of the MN while it locates in its *Home Network*. Care-of Address is an IP address assigned in another network, when the MN is apart from its Home Network.

2.1.2 Rendezvous Protocols

Mobility also requires some means for *rendezvous*, namely creation of a new session with an MN whose current location is unknown.

Rendezvous in Mobile IPv6 One solution is to use a middlebox to forward packets destined for the MN. In Mobile IP, a Home Network has a node called *Home Agent* (HA). An HA always locates in its Home Network and manages identifier-locator bindings of the MNs belonging to the Home Network. When an MN leaves the Home Network and joins in a remote network, the HA of the Home Network behaves like a middlebox between the MN and CNs. After the HA receives a binding update message from the MN, every packet destined for the MN's Home Address is forwarded to the corresponding Care-of Address.

Mobile IP specifies two modes. In bidirectional tunneling mode, packets from an MN to the CN is also forwarded by the MN's HA. Mobile IP also defines route optimization mode, where the MN can notify CNs its own Care-of Address and then take a direct communication in order to avoid a redundant routing.

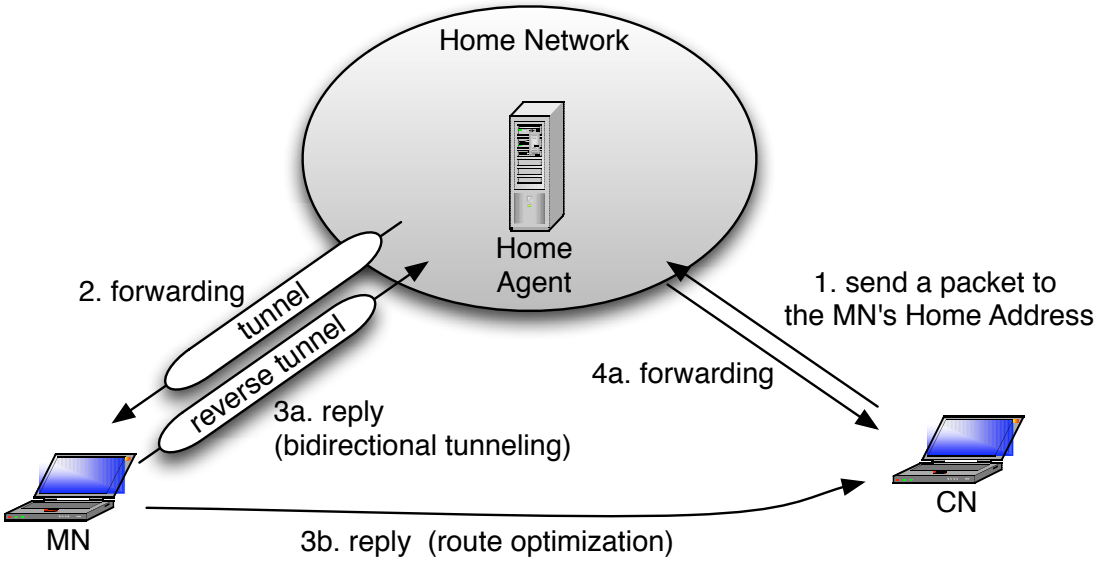

Figure 1: Mobile IPv6

Rendezvous in LIN6 Another solution for rendezvous is to use a query-based system like DNS, though DNS is not suitable due to its slow update mechanism. Each MN stores its identifier-locator binding in a database node in a public network, and maintains the binding. Those who wants to send a packet to an MN can ask the database of the MN for its current locator. For instance, LIN6 follows this manner; *Mapping Agent* works as a public database node and manages bindings between LIN6 ID and IP address (Figure 2).

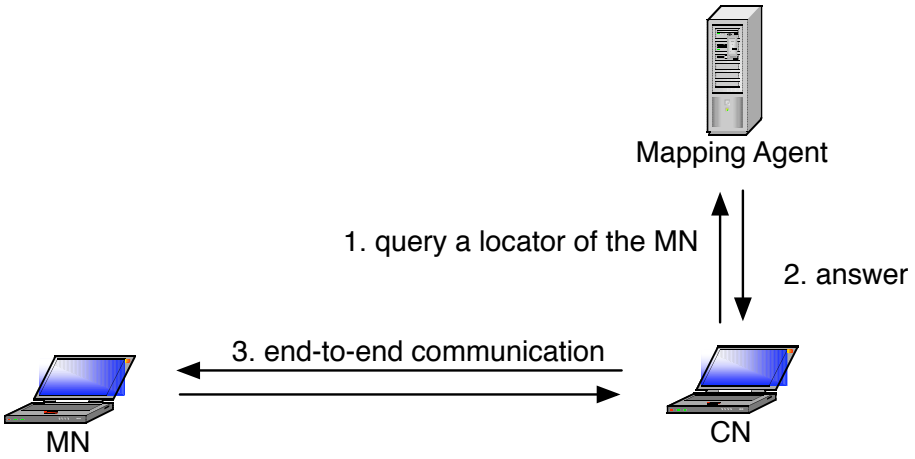

Figure 2: LIN6 communication with Mapping Agent

Anyway, we need a help of another functional entity for a rendezvous mechanism.

2.2 Location Privacy

Location privacy is ability to prevent other parties from learning one's current and/or past location [5]. It concerns *unlinkability* between identifiers when a movement happens. To guarantee complete location privacy, any relation between a location and identifiers of the node must be concealed from others.

2.2.1 Privacy Threat

In general, privacy problems range over multiple layers [6] [7]. It is not solved independently in each layer. There are many types of identifiers, e.g. MAC address in the link layer, IP Address in the network layer, SIP URI or E-mail Address in the application layer, and so on. Some are used globally and others

locally. Some are used permanently and others just temporarily. Sequence Number and Port Number in the transport layer, or SPI value in IPsec are examples of temporary identifiers. End-point identifier is global and constant in natural. Therefore we should carefully consider the unlinkability of location with it.

When an identifier changes and another identifier remains the same, the change can be traced by a node which observed the communication. An attacker might bind all identifiers of a node so that he obtains a complete history of IP address change of the target.

2.2.2 Privacy Goals

Our goal is to get an environment where location privacy is completely protected. All the users are not revealed their location at all in such an environment. Therefore, any identifier of a node must be unlinkable to its locator.

There are potential adversary nodes both in CNs and onlookers. Our primary goal is to conceal roaming from general onlookers. Then, the next priority is to conceal from the CN.

Moreover, we often assume trusted third parties. The location privacy requirement also holds true for them. The word 'trusted' means that they are supposed not to abuse private information of the users. However, the information leakage to the trusted third party by itself should be avoided if possible; Suppose that ISPs or cell phone companies have all your location history over your whole life.

There is a sensitive problem in relation to criminal investigation. Anonymity sometimes hinders police from identifying and locating criminals. In our situation, identification is not interfered. We assume that both end-points authenticate each other using their signature. However, location privacy protection has an influence on difficulty in locating a node. We should not encourage a criminal use.

2.2.3 Location Privacy in Mobile IPv6

Location privacy problems in Mobile IPv6 are summarized in [8]. They consists of two problems:

- To disclose Care-of Address to CNs

- To reveal Home Address to onlookers

Note that a CN ordinarily knows the MN's identifier, namely Home Address, while an onlooker does not know the binding of the MN unless he/she analyzes the header options or encapsulated payload in the flowing packets.

In the route optimization mode, an MN can communicates with its CN through the HA or directly. The MN will choose more preferable way according to who the CN is. As long as the MN use reverse tunneling to the HA, the CN does not know the MN's Care-of Address. By contrast, when the MN takes advantage of route optimization to enjoy a more efficient communication, the CN will learn both HoA and CoA of the MN. When the roaming of the MN happens, the CN in a direct communication will receive the Binding Update message and recognize the roaming. On the other hand, onlookers on the MN-CN path can learn the HoA and CoA of the MN by checking Home Address Option and Routing Header in the packets.

In the bidirectional tunneling mode, it seems to the CN and the onlookers on the HA-CN path, as if the MN locates in its Home Network. Meanwhile, those on the MN-HA path possibly learn the MN's location. It depends on whether the payload in the reverse tunnel is encrypted or not. If the inner IP packet is encrypted in IPsec ESP, the onlookers on the MN-HA path can't distinguish it from ordinary IPsec packet from the CoA to the HoA.

In whichever mode, the HA holds the MN's location all the time. As described in the previous section, the HA is trusted but we should also take this node into consideration.

Of course, there is a trade-off between location privacy and efficiency in the communication. Bidirectional tunneling mode with encryption is not always the best choice.

2.3 Host Identity Protocol (HIP)

Host Identity Protocol (HIP)[4][9][10] supports mobility, multihoming, and security in an integrated fashion. Public/private key pairs are used as end-point identifiers to separate IP address from ID use and also to provide authentication mechanism based on public key cryptography. HIP conceptually introduces a

new layer between network and transport, called Host Identity layer. ID-locator bindings are manged and translated in this layer.

This protocol is still an experimental one and being discussed in the IETF.

2.3.1 Host Identity

Host Identity (HI) is created based on digital signature. Each host owns a public/private key pair, and hosts are identified by the possession of a private key corresponding to particular signature.

Since the size of an HI is not uniquely specified and usually it is quite large (512, 1024, or 2048 bits), *Host Identity Tag* (HIT) is used as an end-point identifier. HIT is a 128-bit value which contains the hash value of the original HI.

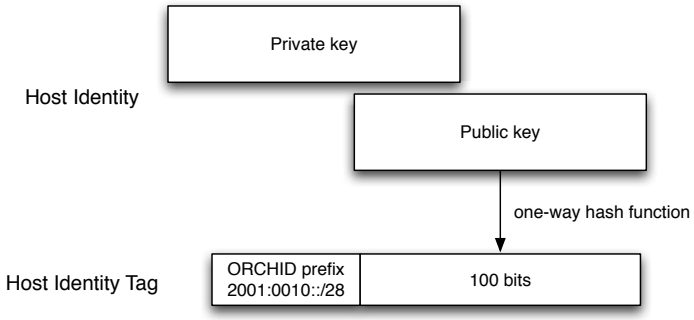

Figure 3: Structure of Host Identity Tag

Figure 3 shows the structure of a HIT. A HIT consists of 28-bit ORCHID prefix and 100-bit hash value. HIT is designed as an ORCHID (Overlay Routable Cryptographic Hash Identifiers) [11], which represents a special class of IPv6 Address. A HIT can be used as an "IPv6 Address" in the application layer. This encourages the reuse of legacy applications.

2.3.2 HIP Base Exchange

In HIP, transport communications are encapsulated with IPsec ESP. Owing to this, end-points have a key exchange phase in the beginning of a HIP session, as in a usual IPsec session. HIP defines a specific key exchange protocol called the HIP base exchange. In this process, the initiator and the responder exchange their HIs as well as the keying material to establish a Security Association (SA).

Figure 4 shows how the base exchange proceeds. The base exchange is a 4-

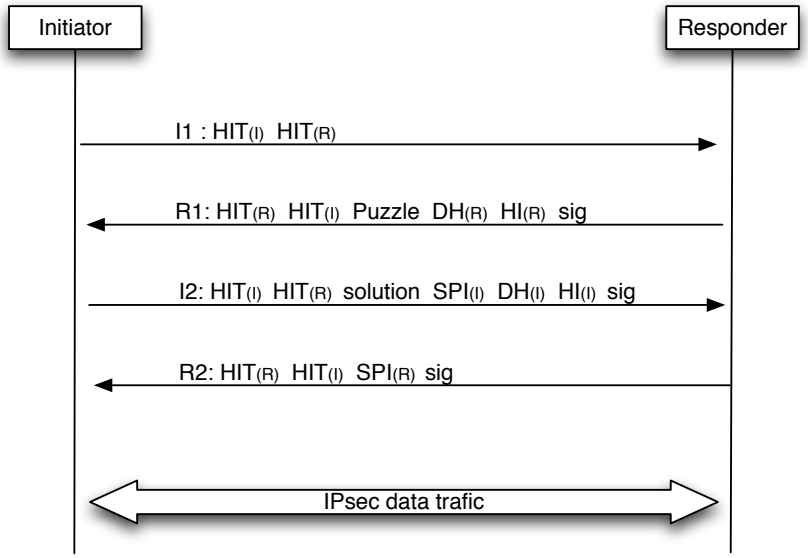

Figure 4: The HIP base exchange protocol

way handshake process, consisting of respectively I1, R1, I2, R2 packets. These packets commonly contains source and destination HITs in their HIP header.

The initiator sends the I1 packet to initiate a base exchange. On receiving the I1 packet, the responder immediately replies a prepared R1 packet. The R1 packet contains a puzzle. The puzzle is a cryptographic question, which ask the initiator to calculate an original number of a hash value by brute-force computation. Usually it takes several seconds. And then, the initiator sends the I2 packet containing the solution of the puzzle. Until the responder checks that the solution is correct, it does not store any state to avoid DoS attacks. In addition, the R1 and I2 packets carry their Host Identity to verify the signature, and a Diffie-Hellman value to calculate a keying material for Security Associations (SAs).

In this process, the end-points create SAs and exchange the SPI values by I2 and R2. After the base exchange, transport communications starts over the established SAs. The HITs are associated with the SPI values, and so they are resolved in the HIP layer.

2.3.3 Rendezvous Mechanism and Mobility

Rendezvous Server (RVS) is a HIP node which performs the I1 packet forwarding [12]. An MN registers in a RVS beforehand, following the HIP registration protocol [13]. Then the CN is able to utilize the RVS to initiate a communication with the MN (Figure 5).

The CN, as an initiator, gets the HIT of the MN and the IP address of the RVS the MN registered with, for example, by DNS. The I1 packet sent by the CN is attached FROM option at the RVS, and forwarded to the MN. Since the MN sees the IP address from the FROM option, it can send the reply (R1) directly to the initiator. The R1 packet includes VIA option so that the initiator can confirm that the packet was forwarded using RVS. The following communication is performed directly between the initiator and the responder.

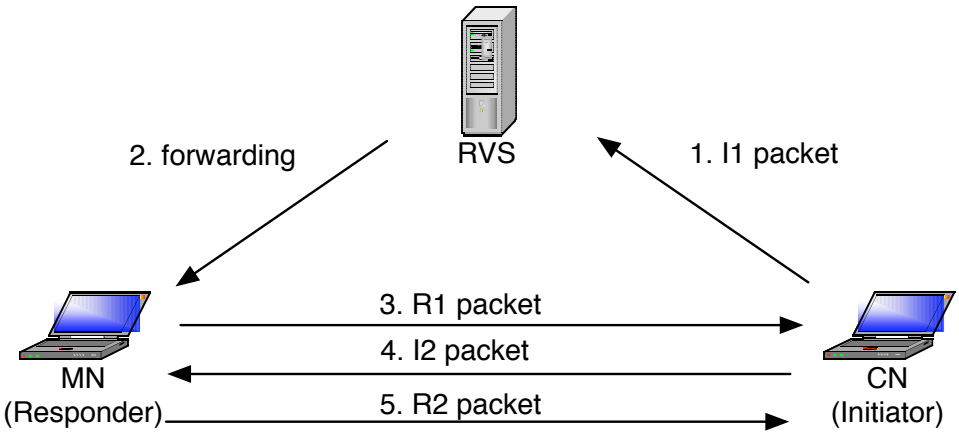

Figure 5: HIP Rendezvous Mechanism

As for the roaming, [14] gives the specification. When a MN change its IP address during a session, the MN sends a HIP UPDATE message to the CN and its RVS, followed by the 3-way messages to confirmation. This process is transparent to the transport layer, similarly to the general mobility protocol.

Moreover, RVS gives a solution for "double jump". When both end-points roam at once and miss the location of the correspondent, they attempt to make a new session through RVS.

2.3.4 Location Privacy in HIP

HIP itself does not have a particular function for location privacy. On the contrary, as usual in mobility protocols, HIP introduces privacy risks with use of a global persistent identifier, which strongly identifies a user. As described in 2.2.1, the user can be tracked by constant identifiers before and after roaming, over multiple protocol layers.

HIP with Upper Layers HIP supports IPsec ESP by default. If the transport data is encrypted by ESP in tunneling mode or BEET[1)] mode, all the header information in transport or upper layers is concealed.

HIP with Lower Layers A natural approach to reduce privacy threats is to change identifiers regularly in a short term. Moreover, identifiers in all the protocol layers should be changed simultaneously.

A HIP privacy architecture in which identifiers in MAC, IP, and HIP/IPsec layers are changed simultaneously was proposed [15]. In this paper, we assume a similar architecture, and do not deal with problems of the lower layers.

[1)] Bound End-to-End Tunnel

Chapter 3 Related Works

In this chapter, we show two location privacy frameworks: HIP Location Privacy Framework and BLIND.

3.1 HIP Location Privacy Framework

This is a framework proposed by Matos et al. [16] and discussed in the Internet-Draft [17].

A special HIP node called *Rendezvous Agent* (RVA) is introduced. An RVA keeps a lot of global IP addresses and lease them to registered MNs. Thus, an registered MN is leased a virtual network interface and multihomed.

There is a local network area which accompanies an RVA, called *RVA protected area*. An RVA is a middlebox which locates at the border of a public and a protected network. It forwards packets in both direction. Traffic in a protected area is separated from public. In this way, An RVA hide the actual locator of an MN from outer nodes.

RVAs also support two kinds of handovers. The intra-RVA handover helps local mobility in a protected area. An MN send UPDATE message to the RVA, so that it maintains its forwarding table. The handover events are not observed by nodes outside.

The inter-RVA handover helps a network migration to another RVA protected area. Suppose that an MN is located at the RVA_1's protected area, and wants to migrate to the RVA_2's. The MN notifies its migration to the RVA_1 and register to the RVA_2. Then, the packets for the MN which has been arrived at RVA_1 after its leaving are forwarded to RVA_2, and finally sent to the MN. Outer nodes can observe an inter-RVA migration, and they get the location of the RVA hiding the MN. However, they cannot learn the exact location of the MN.

3.2 BLIND

The BLIND Framework [18] offers a complete end-to-end identity protection in a static communication. Using a sophisticated key exchange protocol, end-point

identifiers are anonymized to onlookers. Moreover, locators can be concealed from the correspondents by introducing Forwarding Agent (FA).

This framework works with any protocol like HIP, where public keys are used as end-point identifiers. In this section, a HIP specific version of the BLIND is explained.

3.2.1 Blinded Identifiers

In BLIND, raw end-point identifiers are not disclosed in the packet header. *Blinded identifiers* are used instead.

Blinding is an operation defined in this framework, which consists of concatenation and SHA1 hash function. A blinded HIT (BHIT) is generated per key exchange. A different nonce (N_k) is randomly chosen in each time, thus k-th blinding function is defined as follows:

$$\begin{aligned} \text{blind}_k : \text{HIT} &\rightarrow \text{BlindedHIT} \\ p \mid\mid x &\mapsto p \mid\mid \text{hash}(N_k \mid\mid x) \end{aligned}$$

Here, p represents the 28-bit ORCHID prefix and the symbol $||$ represents the bit-concatenation operator. In the k-th session, a HIT h is mapped to $\text{blind}_k(h)$.

Note that end-points are still identified by HIT in upper layers. Mappings between HIT and BHIT are managed in the HIP layer.

3.2.2 Key Exchange Protocol

Figure 6 shows the extended HIP base exchange, with header information in each packet. By comparing with the original HIP base exchange, you will notice that source and destination HITs are replaced by blinded HITs.

When the responder receives an I1 packet, it is seemingly very difficult to determine HIT_R from BHIT_R because it calculates the inverse of one-way hash function. However, if the responder knows all of the candidates for HIT_R and they are few enough, testing each HIT with the given nonce will tell the correct answer at a slight cost. Since it can be considerable cost for a busy server which possesses a lot of HITs, HINT option is defined to reveal a few bits of the raw HIT. The initiator can control the number of bits to be revealed.

Receiving the R1 packet, the initiator can calculate the Diffie-Hellman keying material and immediately encrypt HI_I. The signature in the R1 will be

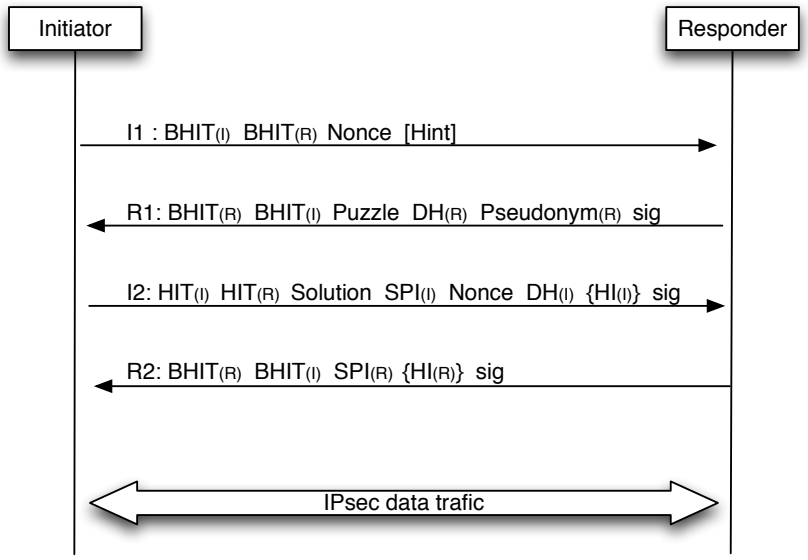

Figure 6: The Blind Base Exchange

verified later, when the initiator receives HI_R in the R2 packet.

3.2.3 Forwarding Agent (FA)

Forwarding Agents (FA) [9] behave like SPINAT[19]. They forward an IPsec packet according to the source/destination addresses and the SPI value in a packet. However, they are different from SPINAT, in that they do not have to be located at the border between public/private networks.

FAs has a number of IP addresses and a HIP node can lease a virtual interface from an FA to be virtually located there. When HIP nodes perform a base exchange through an FA, the FA automatically saves the IP addresses, HITs and SPI values of the both end-points. Similarly, it updates the SPI values in the database when forwarding an UPDATE packet.

In a BLIND context, an initiator obtains complete location privacy. All the other nodes cannot get both of the HIT and IP address of it; Responders cannot get the initiator's IP address, while FAs and other onlookers cannot get the initiator's HIT.

FAs can be trustworthy or untrustworthy. Trustworthy FAs authenticate a host before leasing their interface, while untrustworthy ones allow anonymous leases. When using a trustworthy FA, the user's identity is revealed to the FA, so the user's location privacy is not protected from it. On the other hand, an

untrustworthy FA provides a complete location privacy for users.

However, both nodes rely on the conscience of users. They are vulnerable to attacks of too many leasing by a malicious node. This problem is discussed again in the next chapter.

3.3 Degree of Location Privacy

In this section, we summarize the result of each protocol. They are evaluated in terms of location privacy.

The symbol ◯ shows that the node does not threaten location privacy of the MN, while the symbol × shows that the node can get both identifier and locator of the MN.

3.3.1 Mobile IPv6 (Bidirectional Tunneling)

Table 1 shows the location privacy of an MN in the bidirectional tunneling mode of Mobile IPv6. There are two kinds of routing paths:

- MN → HA → CN
- CN → HA → MN

Table 1: Mobile IPv6 (Bidirectional Tunneling)

Encapsulation	MN-HA	HA	HA-CN	CN
plain	×	×	◯	◯
encrypted	◯			

Because the packets on both directions go through HA, the CN and onlookers on the CN's side do not get the CoA of the MN. On the other hand, onlookers on the MN's side depend on the encryption of the encapsulated packet. The HA always holds the binding of the MN.

3.3.2 HIP Location Privacy Framework

Table 2 shows the location privacy of an MN in HLPF. There are three kinds of paths:

- MN → RVA → CN
- CN → RVS → RVA → MN

- CN → RVA → MN

Table 2: HIP Location Privacy Framework

MN-RVA	RVA	RVA-RVS	RVS	RVS-CN	CN	CN-RVA
–	×	◯	◯	◯	◯	◯

It is similar as Mobile IPv6 in terms of location privacy. The nodes *outside* of the RVA cannot get the real location of the MN. As for the *inner* nodes, it depends on how packets are carried in the RVA protected area and it is not specified. In the case of normal IP, the location is not concealed. The RVA always knows the HI and the IP address of its registrants.

3.3.3 BLIND (direct)

Table 3 shows the location privacy of an MN in BLIND, directly communicating with a CN. There are two kinds of paths:

- MN → CN
- CN → MN

Table 3: BLIND

MN-CN	CN
◯	×

Onlookers on the MN-CN will not get the identifier of the MN. They see only the BHIT. In this case, the MN can roam around networks.

However, the CN can get both HI and IP address of the MN.

3.3.4 BLIND (with FA)

Table 4 shows the location privacy of an MN in BLIND, communicating through an FA. There are two kinds of paths:

- MN → FA → CN
- CN → FA → MN

In trustworthy FA, MNs are authenticated in some way (not limited in Host Identity), therefore the FA knows an identity of the MN and its location. An

Table 4: BLIND with FA

FA type	MN-FA	FA	FA-CN	CN
trustworthy	◯	×	◯	◯
untrustworthy		◯		

untrustworthy FA knows about the MN as well as onlookers.

Anyway, mobility mechanism is not supported in this case.

Chapter 4 Protocol Design

In the BLIND Framework with an FA, an MN can conceal its location from the other nodes. In this chapter, we define a mobility management protocol so as to add in mobility and security to the BLIND Framework.

4.1 Packet Forwarding at FA

4.1.1 Temporary Host Identity (THI)

We introduce Temporary Host Identity (THI) to decouple IDs for communications and those for roaming. THI is an HI of the MN, which is used for FAs to identify the MN. It is an identifier for mobile functionality. General applications do not use it for an end-point identifier. Using THI, the MN can roam around without leaking its global HIT.

4.1.2 THI Registration

First, an MN registers to an FA with its THI. It follows HIP Registration Extension [13].

Later in Section 4.4, we describe an anonymous authentication in the registration phase, where the FA can confirm that the MN is permitted to use the FA without knowing the real identity of the MN.

It is assumed that the FA has quite a large number of IPv6 addresses for a lease. On registration, the FA assigns one of its global IP address to the MN in some finite period of time. As a result, the MN virtually locates at the FA's network.

In this step, the FA stores a 5-tuple in the database (THIDB):

$$(\text{THIT},\ \text{THI},\ \text{IP}_{\text{real}},\ \text{IP}_{\text{forward}},\ \text{Lifetime})$$

THIT (Temporary Host Identity Tag) is a tag for a THI, as a HIT is for an HI. $\text{IP}_{\text{forward}}$ is the FA's address leased to the MN.

4.1.3 Forwarding IP Packets to the MN

When an FA receives a packet and the destination IP address is leased to a registered MN, the packet is forwarded to the MN unless the lease is expired.

Since this might cause some security problems such as DoS attacks, it is

preferable that a user of the MN can control the forwarding rule in some way, e.g. he/she can limit source IP addresses.

4.1.4 Forwarding HIP Packets from the MN

FAs also forward packets sent by a registered MN. This must be done carefully since the destination is arbitrary. It is likely that an MN is spoofed and utilized for attacks.

FAs are designed to forward only valid HIP packets. A HIP session is verified at the FA in the base exchange or the update processes in HIP, checking that the source THIT is registered and the packet is signed by the THI. Following data packets are checked whether the tuple of IP addresses and SPI is belongs to some verified session.

Forwarding Control Packets When an MN sends a HIP control packet [1)] through an FA, a THIT and a signature by THI are attached to the HIP header. The MN designates the destination by the FORWARD_TO option including the IP address of the CN. The FA verifies the signature and makes sure that the sender is registered.

There is a location privacy problem concerned with UPDATE packets. We discuss it in 4.3.3.

Forwarding Data Packets An FA has to maintain the database (SPIDB) to forward data packets encrypted in IPsec. Therefore, the FA monitors base exchanges and UPDATE messages between an MN and a CN, and learns the SPI values corresponding to a BHIT pair.

The SPI database contains 4-tuple entries as follows:

$$(\text{SPI},\ \text{IP}_{\text{src}},\ \text{IP}_{\text{dst}},\ \text{Lifetime})$$

4.2 Rendezvous Support

4.2.1 Rendezvous Request from CN to MN

Since the MN stores in the RVS its FA's locator ($\text{IP}_{\text{forward}}$) as the current address, the Rendezvous process goes well as usual. The RVS will forward the I1 packet to the FA, and then the FA will forward it to the right MN (Figure 7).

[1)] 9 types are defined: I1, I2, R1, R2, UPDATE, NOTIFY, CLOSE, CLOSE_ACK

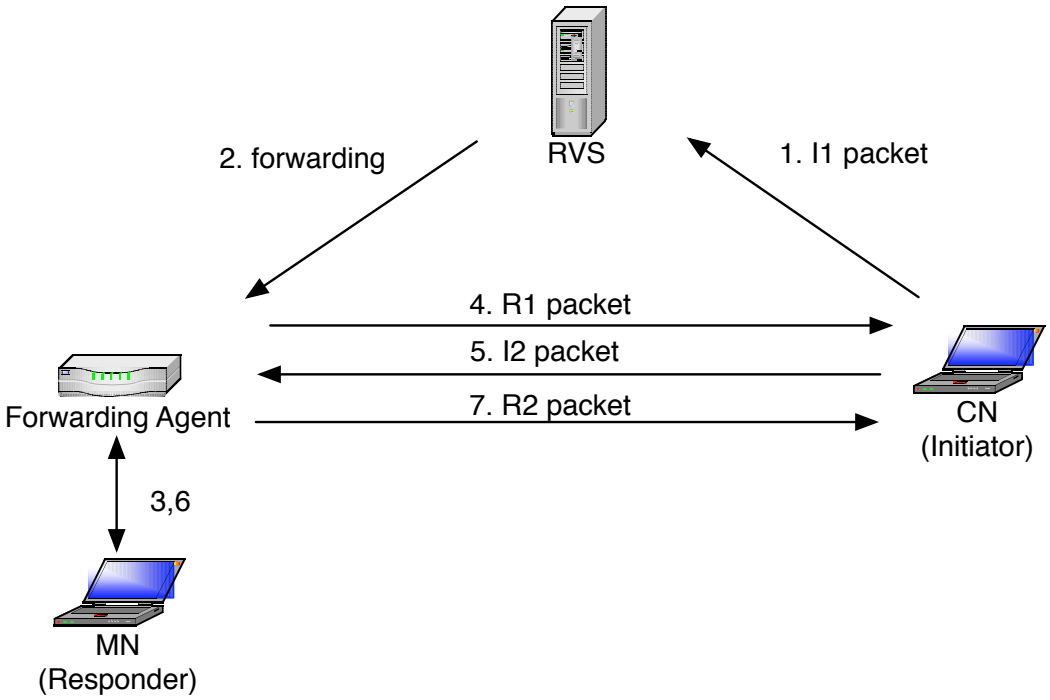

Figure 7: Rendezvous Request from CN to MN

4.2.2 Rendezvous Request from MN to CN

An RVS cannot forward the I1 packets of the blind base exchange since BHITs cannot be identified. In addition, it is not preferable that HIP headers include a raw HIT of a CN, as it indicates that the MN is the one who communicated with a particular CN in a particular time of day.

Therefore, we adopt a HIP-over-HIP encapsulation. Figure 8 shows how the I1 packet to the RVS is carried through IPsec tunnel.

When there are already SAs between the MN and the RVS, the tunnel is reused. Otherwise, the MN will establish new SAs with the RVS, using the blind base exchange.

We expect the efficiency does not fall off so much. Only one packet, I1, is carried in this way in a HIP session. When the MN frequently uses a particular RVS, the HIP session between the MN and the RVS will be kept alive.

4.3 Mobility Support

In this section, the mobility extension to the BLIND Framework is specified.

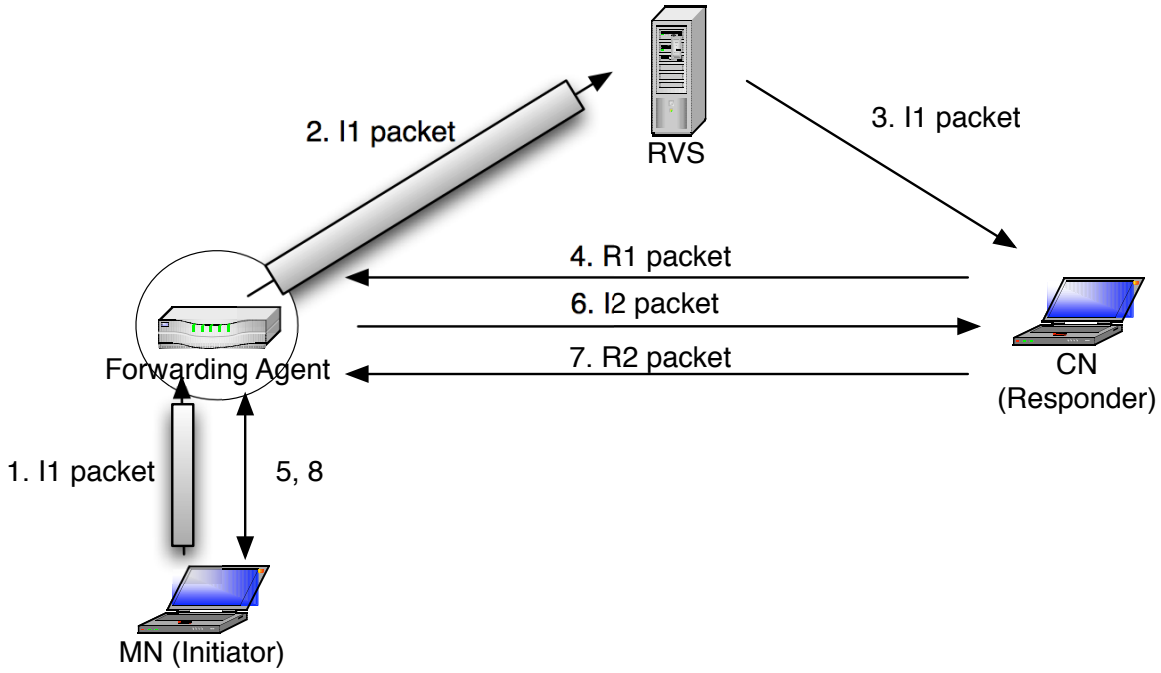

Figure 8: Rendezvous Request from MN to CN

We define two types of roaming. Both do not break an end-point session. The situations are different according to whether or not the MN simultaneously changes the FA.

We refer to the movement without changing FA as a *local jump* (Figure 9), otherwise a *global jump* (Figure 10). By and large, the former is good for hiding from the CN, while the latter for hiding from the FA and onlookers.

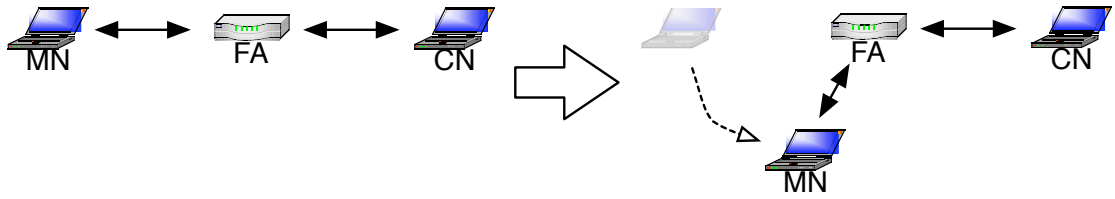

Figure 9: Local Jump

4.3.1 Local Jump

The local jump is specified as follows:

1. The MN gets a new IP address.
2. The MN sends an UPDATE message to the FA, using the registered THI.

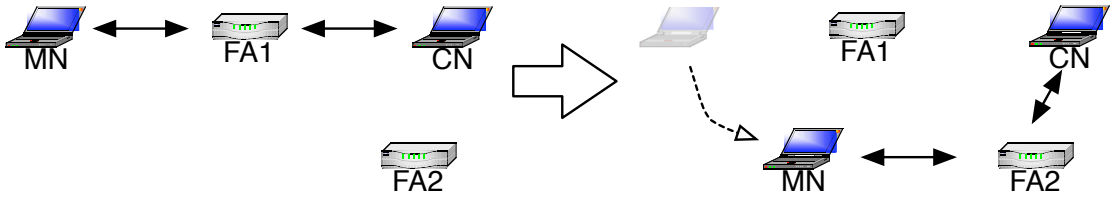

Figure 10: Global Jump

3. HIP mobility update signaling is exchanged between the MN and the FA.
4. At the FA, the IP_{real} field of the MN is updated in the THIDB.

In this session, the CN does not find any indication of roaming from receiving packets, though round trip time might get faster or slower due to the roaming. Meanwhile, the FA and those on the MN-FA path can notice the roaming, tracing the constant THI or SPI values. Nevertheless, they can notice at most that "some one has got moved" since they do not have the actual identity of the MN.

4.3.2 Global Jump

The global jump is specified as follows:

1. The MN gets a new IP address.
2. The MN registers to another FA with a new THI.
3. The MN sends an UPDATE message to the RVS and the CN, maybe after the blind base exchange with them through the new FA.

On the last step, it is recommended to have a rekeying of the Security Associations. Also, the MAC Address of the MN should be changed on the first step. Similarly, all the identifiers of the MN should be changed simultaneously, except the HIs used for the communications with CN. The HIs used in the BLIND framework are concealed from all other nodes. Therefore, the roaming is not detected by the nodes other than the CN.

As for the CN, it will notice the change in FAs. From the viewpoint of the CN, it looks as if the MN had been located at the point of the old FA and had moved to the point of the new FA. The MN can change its FA without roaming as well, whenever it likes.

The cost of registering to a new FA is relatively large, thus packet loss easily happens. It expects retransmission mechanism in upper layers. When the MN

does not have any ongoing session, this mode is recommended.

4.3.3 To Avoid the Location Privacy Risk on UPDATE

Problem As specified in the RFC5202[20], UPDATE packets contain ESP_INFO parameter to tell SPI values used in the session. It indicates an old SPI value which was used in the previous session. In addition, UPDATE packets are supposed not to be encrypted so that smart middleboxes, e.g. SPINAT[19], on the path can monitor the SPI values to maintain a forwarding table. Since our FAs perform a SPI-based packet forwarding, this is also true in our framework.

For this reason, the roaming might be traced by combining the SPI value in the ESP_INFO parameter and the one which had been used before roaming.

An ESP_INFO parameter in a HIP header contains an old SPI value and a new SPI value fields. Combination of the old/new values determines the behavior of rekeying following Table 5.

Table 5: Relation between ESP_INFO and Rekeying Behavior

	old SPI value	new SPI value
no rekeying	an existing value	the same value
rekeying	an existing value	a different non-zero value
new SA	zero	a non-zero value
deprecating the SA	an existing value	zero

Solution We use two UPDATE messages to perform one update. The first UPDATE message adds a new SA to the session. When this packet is forwarded at the FA, a 4-tuple entry is properly created in the SPIDB. At this point, the HIP session is multihomed. An old address with an old SPI remains.

The second UPDATE message depreciates the old SA. This packet includes the old SPI value, and it might reveal the previous location unless it is encrypted. Therefore, we send it as an encrypted HIP data packet. Encryption does not cause a NAT traversal problem since the old SPI value is already obsolete, thus middleboxes do not have to take care of this UPDATE packet. The old SPI value will be expired in the old FA after a while.

4.4 Authorized Temporary Host Identity (ATHI)

As described in 3.2.3, unconditional forwarding service for any Temporary Host Identity might cause serious attacks to the FA. An FA has to accept all anonymous nodes for only a *temporary* identity. A malicious node will register to an FA with thousands of THIs. The identifiers are self-generated public keys. Thus the FA cannot distinguish malicious nodes from other ordinary users. This problem can be also applied to the original BLIND Framework.

Therefore, we need a way to authenticate users without knowing their real identity. The solution we suggest is the authorization using blind signature. We assume there the FA can authenticate the registered users in some way, for example, using the user's HI.

4.4.1 Authorization Using Blind Signature

Blind Signature is a kind of digital signature protocols, introduced by Chaum [21][22]. It is a bit misleading but the word 'blind' has no relationship with the BLIND Framework. It is a protocol for a situation where Alice wants to make Bob sign to her secret text, however, she does not want Bob to read the content.

1. The MN generates a THI
2. The MN ask the FA of sign to the THI
3. The FA authenticates the MN based on the real identity
4. The FA signs the THI by Blind Signature
5. When using THI, the MN shows the FA's signature to the FA

The signature issued by the FA should have a time limit. Otherwise, the size of table will continue to increase.

4.4.2 Other Problems

Restricting users leads to reveal some information about the user. If very few nodes registered with an FA, a use of the FA has a big deal of information. Therefore, FAs which is used for location privacy protection should hold enough number of users all the time, though this would be not easy.

Chapter 5 Evaluation

5.1 Achievements

- MN → FA → CN
- CN → FA → RVS → MN
- CN → FA → MN

Table 6: Location Privacy Protection of Our Framework

CN	FA	RVS	onlookers
○	○	○	○

Since our protocol is based on the BLIND with untrustworthy Forwarding Agent, the MN obtains complete location privacy. Also, our protocol provides mobility.

5.2 Incident Scenario

Now we take a look at the incidents involved in our framework. We clarify the obligation of each node.

5.2.1 Collusion

The CN and the FA can collude to compromise the MN's location privacy. If the CN gives the raw HI of the MN and the corresponding BHIT, while the FA gives the IP address of the MN corresponding to the given BHIT, then they obtain both HI and IP address of the MN.

5.2.2 Adversary MN

If the MN is a malicious node and took an active attack to the CN, then police will investigate the CN's and FA's logs, so as to identify, locate, and arrest the MN.

For this purpose, end-points both should store logs of HIT-BHIT pair, and FAs should store logs of IP-BHIT pair.

5.3 Fault Tolerance

In our framework, there are two types of nodes as single points of failure, namely FA and RVS. When one of these nodes gets hung-up suddenly, the communication in progress will be stopped.

Multiplexing these nodes is an approach in case of this kind of failure.

5.3.1 Multiplexing Forwarding Agent

Since HIP is equipped with multihoming, we can utilize it for multiplexing FA and RVS. If the MN is multihomed, the failure will be automatically recovered.

The MN has a little drawback for multihoming in terms of location privacy. The Security Associations of the MN are possibly related due to the same IP address, though its identifiers are concealed as well.

5.3.2 Multiplexing Rendezvous Server

Similarly, the MN can also register to some RVSs at once. Packets are correctly forwarded regardless of the RVSs.

Besides, RVS can be operated in a distributed environment such as Distributed Hash Table (DHT), where the nodes in disorder are automatically avoided. In such case, the MN does not have to care about a failure in the RVS to some degree.

5.4 Scalability

Currently, we do not specify any rules in selecting an FA. It should be controlled based on the privacy policy of each node. However, it might cause access convergence in particular FAs, and this can be an obstacle for scalability.

One way of avoiding congestion is to introduce *area*. It can be used to decrease concentration of users, for example, distant people are not permitted to use the FA. However, it also might cause privacy leakage, and we should careful for that.

Probably we should introduce a congestion control system, so that clients can select an FA based on topological distance and traffic.

5.5 Service Model

In this section, we discuss a service model of FAs. As we saw in the previous chapter, FAs are placed to serve identifiable users.

Suppose that a forwarding service provider manages a lot of FAs all over the world. A user applies for the forwarding service, and he registers his public key for authentication. Now, the user generates a set of THIs, then ask the provider to sign his THIs. After authentication, he will get THIs with signature by the blind signature protocol. He has a list of the FAs, and choose one. Using one of the THIs with signature, he can register to the FA and take advantage of it.

Chapter 6 Conclusion

In this paper, we discussed IP mobility protocols and the location privacy problem. We proposed a new framework using HIP, and we showed that it is possible to protect location privacy from all other nodes in IP communication with mobility. We defined an extensional mobility management protocol of BLIND, based on a mechanism to separate IDs for mobility from those for end-to-end communication. Also we constructed a service registration with anonymous authentication using blind signature. We evaluated our framework by comparing with existing researches and considering, fault tolerance, scalability and service model.

A congestion controlling discussed in 5.4 is a future work. Some experiments and evidence that it will work in practice are also desired to be done.

Acknowledgments

I would like to express my gratitude to Prof. Yasuo Okabe for giving me a great help, fine suggestions, and all the educational comments for me. I would give my thankfulness to Assoc Prof. Hiroki Takakura and Prof. Motonori Nakamura for significantly helpful discussions.

I really thank all the members of Okabe Laboratory for helping my research. Especially I am grateful to Mr. Jungsuk Song and Mr. Koji Kobayashi for good advice on composing a paper, and my friend, Keita Shimizu for always amusing and encouraging me in daily life.

Finally, I would like to thank my family for supporting me for a long while.

References

[1] Perkins, C.: IP Mobility Support for IPv4, RFC 3344 (Proposed Standard) (2002). Updated by RFC 4721.

[2] Johnson, D., Perkins, C. and Arkko, J.: Mobility Support in IPv6, RFC 3775 (Proposed Standard) (2004).

[3] Kunishi, M., Ishiyama, M., Uehara, K., Esaki, H. and Teraoka, F.: LIN6: A New Approach to Mobility Support in IPv6, *International Symposium on Wireless Personal Multimedia Communication*, Vol. 455 (2000).

[4] Moskowitz, R., Nikander, P., Jokela, P. and Henderson, T.: Host Identity Protocol, RFC 5201 (Experimental) (2008).

[5] Haddad, W. and Nordmark, E.: Privacy Aspects Terminology, Internet Draft (Work in Progress) (2008).

[6] Haddad, W., Nordmark, E., Dupont, F., Bagnulo, M. and Patil, B.: Privacy for Mobile and Multi-homed Nodes: Problem Statement, Internet Draft (Work in Progress) (2006).

[7] Haddad, W., Nordmark, E., Dupont, F., Bagnulo, M., Park, S. S. D., Patil, B. and Tschofenig, H.: Anonymous Identifiers (ALIEN): Privacy Threat Model for Mobile and Multi-Homed Nodes, Internet Draft (Work in Progress) (2006).

[8] Koodli, R.: IP Address Location Privacy and Mobile IPv6: Problem Statement, RFC 4882 (Informational) (2007).

[9] Nikander, P., Ylitalo, J. and Wall, J.: Integrating Security, Mobility and Multi-Homing in a HIP Way, *NDSS* (2003).

[10] Gurtov, A.: *Host Identity Protocol (HIP): Towards the Secure Mobile Internet*, Wiley, New York (2008).

[11] Nikander, P., Laganier, J. and Dupont, F.: An IPv6 Prefix for Overlay Routable Cryptographic Hash Identifiers (ORCHID), RFC 4843 (Experimental) (2007).

[12] Laganier, J. and Eggert, L.: Host Identity Protocol (HIP) Rendezvous Extension, RFC 5204 (Experimental) (2008).

[13] Laganier, J., Koponen, T. and Eggert, L.: Host Identity Protocol (HIP)

Registration Extension, RFC 5203 (Experimental) (2008).

[14] Nikander, P., Henderson, T., Vogt, C. and Arkko, J.: End-Host Mobility and Multihoming with the Host Identity Protocol, RFC 5206 (Experimental) (2008).

[15] Takkinen, L.: Host Identity Protocol Privacy Management, Master's thesis, Helsinki University of Technology (2006).

[16] Matos, A., Santos, J., Sargento, S., Aguiar, R., Girão, J. and Liebsch, M.: HIP location privacy framework, *Proceedings of first ACM/IEEE international workshop on Mobility in the evolving internet architecture*, ACM Press New York, NY, USA, pp. 57–62 (2006).

[17] Matos, A., Santos, J., Girao, J., Liebsch, M. and Aguiar, R.: Host Identity Protocol Location Privacy Extensions, Internet Draft (Work in Progress) (2006).

[18] Ylitalo, J. and Nikander, P.: BLIND: A Complete Identity Protection Framework for End-Points, *Lecture Notes In Computer Science*, Vol. 3957, p. 163 (2006).

[19] Ylitalo, J., Salmela, P. and Tschofenig, H.: SPINAT: Integrating IPsec into Overlay Routing, *International Conference on Security and Privacy for Emerging Areas in Communications Networks (SECURECOMM'05)*, Vol. 0, pp. 315–326 (2005).

[20] Jokela, P., Moskowitz, R. and Nikander, P.: Using the Encapsulating Security Payload (ESP) Transport Format with the Host Identity Protocol (HIP), RFC 5202 (Experimental) (2008).

[21] Chaum, D.: Blind Signatures for Untraceable Payments, *CRYPTO'82*, pp. 199–203 (1982).

[22] Chaum, D.: Blind Signature System, *CRYPTO'83*, p. 153 (1983).